一般の線型回帰モデルの理論

沖 津　　直

1．一般の線型回帰モデルの諸前提

ある変数 y と $k-1$ 個の説明変数 $X_2, X_3, \cdots, X_k$ と攪乱項 u の間に，ある線型関係が存在するものとする。われわれの目前には被説明変数 y と $k-1$ 個の説明変数のそれぞれ n 個の観測値からなるある任意の標本があるものとする。一般に，母集団の y および $k-1$ 個の変数の線型関係は次のように表現できる。

$$Y_i=\beta_1+\beta_2X_{2i}+\cdots+\beta_kX_{ki}+\mu_i \qquad i=1, 2, \cdots, n \tag{1}$$

ここで，係数 β_i は母集団の母数すなわちパラメータ (Parameters) を，また，μ_i は攪乱項 (disturbance term) と呼ばれる確率変数を表すが，β_i および μ_i の確率分布は未知である。まず当面のわれわれの問題は，これらの未知の確率変数の推定量を得ることである。（1）式の方程式は，行列を使って表わすと，次のような簡潔な式で再表現することができる。

$$\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\mathbf{u} \tag{2}$$

$$
\mathbf{y}=\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \mathbf{X}=\begin{pmatrix} 1 & X_{21}\cdots\cdots X_{k1} \\ 1 & X_{22}\cdots\cdots X_{k2} \\ \vdots & \vdots \qquad\quad \vdots \\ 1 & X_{2n}\cdots\cdots X_{kn} \end{pmatrix}
$$

$$
\boldsymbol{\beta}=\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \mathbf{u}=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \tag{3}
$$

ここで$\mathbf{y}$，$\boldsymbol{\beta}$および$\mathbf{u}$は$n\times 1$の列ベクトル，$\mathbf{X}$は，$n\times k$の行列である。$\mathbf{y}$と$\mathbf{X}$は標本の個々の観測値から成り立つ。

行列$\mathbf{X}$の１からなる第１列は，$\mathbf{y}$の切片を表わすし，行列$\mathbf{X}$のX_{ki}は変数X_kの第i番目の観測値を表わす。この行列$\mathbf{X}$は通常の表示法とは，行と列が逆になっている。

係数$\boldsymbol{\beta}$のベクトルの推定を行うにあたって，われわれは，攪乱項$\mathbf{u}$と観測値を要素とする行列に関してのいくつかの基本的仮定を設定しなければならない。これらの仮定は，推定作業に不可決である。もし，これらのうち１つあるいは，それ以上のものが満足されないときには，モデルそのものを部分的に修正していかなければならない。一般の線型回帰モデルの基本的仮定は次のとおりである。

$E(\mathbf{u})=\mathbf{0}$　(4－a)

$E(\mathbf{u}\mathbf{u}')=\sigma^2\mathbf{I}_n$　(4－b)

$\mathbf{X}$は固定的な数の集合である。　(4－c)

$\mathbf{X}$のランク$k<$観測数n　(4－d)

(4－a)の仮定は確率変数である攪乱項u_i($i=1, 2, \cdots, k$)は，確率的な動きをするがその期待値がゼロであること，すなわち，u_iの値はゼロよりも大きくなったり小さくなったりするが平均すれば，ゼロになるという仮定である。また，(4－b)の仮定は攪乱項u_i($i=1, 2, \cdots, k$)の分散が一定であることを表わしているが，$\mathbf{u}$が$n\times 1$のベクトルであるから$\mathbf{u}\mathbf{u}'$はn

次の対称行列となり，$\mathbf{u}$の分散共分散行列は，以下のようになる。

$$E(\mathbf{u}\mathbf{u}') = \begin{bmatrix} E(u_1^2) & E(\mathbf{u}_1 u_2) & \cdots E(u_1 u_n) \\ E(u_2 u_1) & E(u_2^2) & \cdots E(u_2 u_n) \\ \vdots & \vdots & \vdots \\ E(u_n u_1) & E(u_n u_2) & \cdots E(u_n^2) \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$

$E(u_i)^2=\sigma^2$(一定)，すなわちすべてのiについてu_iの分散は一定の値σ^2をもつがそれ以外の項は$E(u_t u_{t+s})=0$，$s\neq 0$，すなわちu_tとu_{t+s}変数は互いに無相関，つまり互いに独立である。

（4－c）の仮定は（k－1）個の説明変数といわれる観測値からなる行列であるが分布を伴なう確率変数ではなく，固定的なものであるということである。つまり$\mathbf{X}$には具体的な値が指定される。$\mathbf{u}$に関する（4－a），（4－b）の仮定を考慮するならば，$\mathbf{y}$の分布は$\mathbf{X}$の値が与えられたときの$\mathbf{y}$の条件付期待値との線型回帰方程式で表わされ，その分布がその超平面上にその平均があり，分散が$\mathbf{u}$と同じく$\sigma^2\mathbf{I}_n$ということである。

最後に（4－d）の仮定はk－1個の説明変数のそれぞれn個の観測値からなる行列のランクが観測値の個数よりも大きくなければならないということである。もし，そうでなければ逆行列が存在しなくなるからである。

以上が母集団に関するデータを要約した諸前提である。多くの場合，われわれの手許にある利用可能なデータは，ほとんど母集団の一部である標本である。その標本は，母集団の特徴を忠実に反映しているものと考えられるから，大切に扱い推測統計学では，それを最も効率的に有効に使って，母集団に関するいろいろな情報を引き出す。母集団の完全な姿を確実に知るには，当然その母集団を構成するすべての要素がわかっていなければならない。しかし，それは，通常無理な要求である。というのは，ある場合には，母集団のすべての要素を調査することがまた，ある場合には，いろいろな現実的な諸制約からたとえば，調査に長時間がかかるとか，莫大な費用がかかるとかの理由からできないからである。このような事情から，標本を利用して母集団を把握しようとする推測統計学が必要となってくるわけである。

2. 最小二乗推定量

われわれが実際に知りたいのは母集団の母数（Parameters）なのであるが，通常それは全数調査でもしない限り，知ることはできない。そこで，それらを知る方法として標本をもとにして，なるべく推定誤差が小さくなるように推定する。

さて，1であげた4つの基本的諸仮定のもとでまず，母集団のパラメータ $\boldsymbol{\beta}$ を推定することからはじめよう。標本の観測値に，攪乱項の推定値の誤差の二乗和が最小になるように回帰平面（超平面）をあてはめる。攪乱項 $\mathbf{u}$ の推定量を $\mathbf{e}$ とし母数の推定量を $\hat{\boldsymbol{\beta}}$ で表わすと，(2)式の母集団の回帰方程式は次のようになる。

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \qquad (5)$$

この推定は，ガウスーマルコフの定理から，最小二乗法の原理を適用して求められている。つまり（5）式の残差項 $\mathbf{e}$ の二乗和が最小になるような線型回帰の標本方程式を求める。すなわち，(5)式の残差項 $\mathbf{e}$ の二乗和は，

$$\begin{aligned}\mathbf{e}'\mathbf{e} &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \ (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{y}'\mathbf{y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\ \mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} \qquad (6) \\ &(\because (\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}) = \mathbf{y}'\mathbf{X}\hat{\boldsymbol{\beta}})\end{aligned}$$

である。この（6）式を列ベクトル $\hat{\boldsymbol{\beta}}$ で偏微分してゼロベクトルとおけばよいから，

$$\frac{\partial}{\partial\hat{\boldsymbol{\beta}}}(\mathbf{e}'\mathbf{e}) = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0}$$

$$\therefore \quad \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$$

$$\therefore \quad \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \qquad (7)$$

（7）式の右辺は，すべて値のわかっている観測値であるから，これを計算することによって $\hat{\boldsymbol{\beta}}$ の値が求められる。すなわち $\hat{\boldsymbol{\beta}}$ は最小二乗推定量である

が，推定量$\hat{\beta}$が望ましい推定量のうち最も重要な不偏性をもっているかを調べるために$\hat{\boldsymbol{\beta}}$の期待値を求めてみよう。（2）式の$\mathbf{y}$を（7）式の$\mathbf{y}$に代入すると，

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta}+\mathbf{u})$$

$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} \qquad (8)$$

$$\therefore E(\hat{\boldsymbol{\beta}}) = E[\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}]$$

$$= E(\boldsymbol{\beta}) + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u})$$

$$= \boldsymbol{\beta} \qquad (9)$$

$$(\because E(\mathbf{u}) = \mathbf{0})$$

つまり標本から求められた最小二乗推定量$\hat{\boldsymbol{\beta}}$が不偏性をもつことが確認された。

次に最小二乗推定量$\hat{\boldsymbol{\beta}}$の分散を求めてみよう。この分散は小さければ小さいほどよい。標本回帰方程式の超平面の傾きのゆれ（Wobble）が少なくてすむからである。（8）式より

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} \qquad (8)'$$

$$\therefore \quad Var(\hat{\boldsymbol{\beta}}) = E[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})']$$

$$= [(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]$$

$$= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \qquad (10)$$

$$\left(\because((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u})' = \mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\right)$$

既に（9）式より最小二乗推定量$\hat{\boldsymbol{\beta}}$は$\boldsymbol{\beta}$の不偏推定量であることが証明されている。ここで，$\hat{\boldsymbol{\beta}}$の分散が他のいかなる推定量よりも小さいことを示すことができれば，最小二乗推定量$\hat{\boldsymbol{\beta}}$は$\boldsymbol{\beta}$の最良線型不偏推定量であることが明らかになる。最小分散性を示すために，要素が既知である$\mathbf{c}$（$k\times 1$の列ベクトル）を考える。$\mathbf{c}'\boldsymbol{\beta}$の可能な推定量の1つが$\mathbf{c}'\hat{\boldsymbol{\beta}}$である。

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \qquad (7)$$

であったから

$$E(\mathbf{c}'\hat{\boldsymbol{\beta}}) = \mathbf{c}'\boldsymbol{\beta}$$

である。したがって分散は，次のようになる。

$$Var(\mathbf{c}'\hat{\boldsymbol{\beta}}) = E[(\mathbf{c}'\hat{\boldsymbol{\beta}}-\mathbf{c}'\boldsymbol{\beta})(\mathbf{c}'\hat{\boldsymbol{\beta}}-\mathbf{c}'\boldsymbol{\beta})']$$
$$= E[\mathbf{c}'(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}')\mathbf{c}]$$
$$= \sigma^2\mathbf{c}'(\mathbf{X}'\mathbf{y})^{-1}\mathbf{c} \tag{11}$$

ここで，他の任意の$\mathbf{c}'\boldsymbol{\beta}$の線型不偏推定量を

$$b = \mathbf{a}'\mathbf{y} \tag{12}$$

と定義してみる。もしこのbが$\mathbf{c}\boldsymbol{\beta}$の不偏推定量であるとすれば，

$$E(b) = E[\mathbf{a}'(\mathbf{X}\boldsymbol{\beta}+\mathbf{u})]$$
$$= \mathbf{a}'\mathbf{X}\boldsymbol{\beta} = \mathbf{c}'\boldsymbol{\beta}$$

でなければならない。すなわち不偏性が成り立つためには

$$\mathbf{a}'\mathbf{X} = \mathbf{c}' \tag{13}$$

でなければならない。すなわち（13）式が成り立つときのみに

$$b = \mathbf{c}'\boldsymbol{\beta} + \mathbf{a}'\mathbf{u}$$

となり，bの分散は

$$Var(b) = E[(b-\mathbf{c}'\boldsymbol{\beta})(b-\mathbf{c}'\boldsymbol{\beta})']$$
$$= E[\mathbf{a}'\mathbf{u}\,\mathbf{u}'\mathbf{a}] = \sigma^2\mathbf{a}'\mathbf{a} \tag{14}$$

となる。また（13）式を使って（11）式を書きかえると

$$Var(\mathbf{c}'\hat{\boldsymbol{\beta}}) = \sigma^2\mathbf{a}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{a}$$

かくて，

$$Var(b) - Var(\mathbf{c}'\hat{\boldsymbol{\beta}})$$
$$= \sigma^2\mathbf{a}'[\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{a}$$
$$= \sigma^2\mathbf{a}'\mathbf{M}\mathbf{a} \tag{15}$$

ここで，

$$\mathbf{M} = \mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \tag{16}$$

$\mathbf{M}$は対称かつ巾等行列である。さらに，それは，すぐ後の（20）式$\mathbf{e}'\mathbf{e} = \mathbf{u}'\mathbf{M}\mathbf{u}$で示されるように，正半定符号 (Positive semi-definite) である。$\mathbf{e}'\mathbf{e} \geqq \mathbf{0}$。従って，$\mathbf{M}$も正半定符号であるから

$$Var(b) - Var(\mathbf{c}'\hat{\boldsymbol{\beta}}) \geqq \mathbf{0} \tag{17}$$

すなわち，$\mathbf{c}'\hat{\boldsymbol{\beta}}$の最小二乗推定量は，他のいかなる線型推定量よりも小さい

分散をもっている。従って，$\mathbf{c}'\hat{\boldsymbol{\beta}}$ は最良線型不偏推定量であることが証明された。そして，

$$\mathbf{c}'_i = [0\cdots0,\ 1,\ 0,\ \cdots0]$$

において i 番目の要素が1で，その他はすべて0である場合が，ちょうどこれまでの推定量に該当することがわかる。

次に残差分散について調べてみよう。

観測値をあてはめた後の行列表示の回帰方程式（5）式より

$$\begin{aligned}\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{u} - \mathbf{X}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \mathbf{u})] \\ &= \mathbf{u} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} \\ &= [\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{u} \\ &= \mathbf{M}\mathbf{u} \qquad (19)\end{aligned}$$

（19）式からわかるように観測される偏差 $\mathbf{e}$ は，未知の攪乱項 $\mathbf{u}$ の線型関数として表現される。従って，残差分散は，

$$\begin{aligned}\mathbf{e}'\mathbf{e} &= \mathbf{u}'\mathbf{M}\mathbf{M}\mathbf{u} \\ &= \mathbf{u}'\mathbf{M}\mathbf{u} \\ &= \mathbf{u}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{u} \qquad (20)\end{aligned}$$

となる。両辺の期待値をとると，

$$\begin{aligned}E(\mathbf{e}'\mathbf{e}) &= E[\mathbf{u}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{u}] \\ &= \sigma^2\{n - tr[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})]\} \\ &= (n-k)\sigma^2 \qquad (21)\end{aligned}$$

ゆえに，σ^2 の不偏推定量 S^2 は次のようになる。

$$\therefore \quad S^2 = \frac{\mathbf{e}'\mathbf{e}}{n-k} \qquad (22)$$

ここで，観測値のデーターに標本回帰方程式をあてはめることによって Y の総変動（$\bar{Y}$ から測って）のうち，どれだけが説明されたかの割合を測る尺度である決定係数（あるいは重相関係数）の式を導出することにする。

まず，Y の総変動 Σy_i^2 および残差分散 Σe_i^2 は，それぞれ

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} Y_i^2 - \frac{1}{n}(\sum_{i=1}^{n} Y_i)^2$$

$$= \mathbf{y}'\mathbf{y} - \frac{1}{n}(\bar{\Sigma} Y)^2 \qquad (23)$$

$$\Sigma e_i^2 = \mathbf{e}'\mathbf{e} - (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

$$(\because \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}(\text{スカラー}) = \mathbf{y}'\mathbf{X}\hat{\boldsymbol{\beta}})$$

$$= \mathbf{y}'\mathbf{y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

$$= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} \qquad (24)$$

である。従って，説明される変動は$\Sigma y_i - \Sigma e_i^2$となる。すなわち，

$$\therefore \quad \Sigma y_i^2 - \Sigma e_i^2 = \mathbf{y}'\mathbf{y} - \frac{1}{n}(\Sigma Y)^2 - (\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y})$$

$$= \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{1}{n}(\Sigma Y)^2 \qquad (25)$$

となる。従って，決定係数$R^2{}_{1\cdot2\cdot3\cdots\cdots k}$は，次のようになる。

$$R^2_{1\cdot2\cdot3\cdots\cdots k} = \frac{\Sigma y_i^2 - \Sigma e_i^2}{\Sigma y_i^2} = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{1}{n}(\Sigma Y)^2}{\mathbf{y}'\mathbf{y} - \frac{1}{n}(\Sigma Y)^8} \qquad (26)$$

ここで，$\frac{1}{n}(\Sigma Y_i^2)$ を$S_{\mathbf{yy}}$，$\frac{1}{n}(\Sigma e_i^2)$ をS_Eと書き表すとすれば (26) 式は，

$$R^2_{1\cdot2\cdot3\cdots\cdots k} = 1 - \frac{S_E}{S\mathbf{y}\mathbf{y}} \qquad (27)$$

ともかける。しかし，この (27) 式の$S_{\mathbf{yy}}$およびS_Eは，それぞれ偏っているために決定係数の大きさは大きめに推定される。

$S_{\mathbf{yy}}$およびS_Eの不偏分散を使った場合の決定係数を調整済決定係数$\bar{R}^2_{1\cdot2\cdot3\cdots\cdots k}$は，次のようになる。

$$\bar{R}^2_{1\cdot2\cdot3\cdots\cdots k} = 1 - \frac{S_E}{S_{\mathbf{yy}}} \cdot \frac{n-1}{n-k}$$

$$= 1 - \frac{n-1}{n-k}(1 - R^2_{1\cdot2\cdot3\cdots\cdots k}) \qquad (28)$$

以上の諸結果を要約して，ここに再掲載しておこう。

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \qquad (2\text{および}5)$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{7}$$

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} \tag{9}$$

$$Var(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \tag{10}$$

$$E(\mathbf{e}'\mathbf{e}) = (n-k)\,\sigma^2 \tag{21}$$

$$R^2 = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{1}{n}(\Sigma Y)^2}{\mathbf{y}'\mathbf{y} - \frac{1}{n}(\Sigma Y)^2} \tag{26}$$

最後に，ここで後述の説明を進めるうえで必要になる偏差の表現形式を展開しておくことにしよう。（1）式および（5）式を標本の大きさ n で除すと，

$$\begin{aligned}\bar{Y} &= \beta_1 + \beta_2\bar{X}_2 + \cdots\cdots + \beta_k\bar{X} + \bar{u} \\ &= \hat{\beta}_1 + \hat{\beta}_2\bar{X}_2 + \cdots\cdots + \hat{\beta}_k\bar{X}_k\end{aligned} \tag{29}$$

（1）式から（29）式を差し引くと，

$$\begin{aligned}Y_i - \bar{Y} &= \beta_2(X_2 - \bar{X}_2) + \cdots\cdots + \beta_k(X_k - \bar{X}_k) + u_i - \bar{u} \\ &= \hat{\beta}_2 x_{2i} + \cdots\cdots + \hat{\beta}_k x_{ki} + e \qquad i = 1, 2, \cdots, n\end{aligned} \tag{30}$$

偏差を小文字で示すと（30）式は

$$\begin{aligned}y_i &= \beta_2 x_{2i} + \cdots\cdots + \beta_k x_{ki} + u_i - \bar{u} \\ &= \hat{\beta}_2 x_{2i} + \cdots\cdots + \hat{\beta}_k x_{ki} + e_i \qquad i = 1, 2, 3, \cdots, n\end{aligned} \tag{31}$$

となる。（31式）を行列表示すると，

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} - \bar{\mathbf{u}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \tag{32}$$

しかし，（32）式の $\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$ は（5）式のそれとは要素が異なっている。

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} x_{21} & x_{31} \cdots\cdots x_{k1} \\ x_{22} & x_{32} \cdots\cdots x_{k2} \\ \vdots & \vdots \qquad \vdots \\ x_{2n} & x_{3n} \cdots\cdots x_{kn} \end{pmatrix} \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_2 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$$

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \quad \bar{\mathbf{u}} = \begin{pmatrix} \bar{u} \\ \bar{u} \\ \vdots \\ \bar{u} \end{pmatrix} \quad \mathbf{e} = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{pmatrix} \tag{33}$$

この（32）式の偏差の形で示される場合においても（7）式の場合と同様に

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

が成立する。また，$E(\hat{\boldsymbol{\beta}})$ および$Var(\hat{\boldsymbol{\beta}})$ を求めてみよう。上の式に（32）式を代入すると，

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \mathbf{u} - \bar{\mathbf{u}}) \\ &= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\bar{\mathbf{u}}\end{aligned}$$

従って，$E(\hat{\boldsymbol{\beta}})$ および$Var(\hat{\boldsymbol{\beta}})$ は次のようになる。

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} \qquad (\because E(\mathbf{u}) = \mathbf{0},\ \mathbf{X}'\bar{\mathbf{u}} = \mathbf{0})$$

$$Var(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$$

また，残差分散の期待値および残差分散もこれまでと同じく次のようになる。

$$E(\mathbf{e}'\mathbf{e}) = (n-k)\ \sigma^2$$

$$\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$$

これまでの場合と異なってくるのは決定係数である。偏差の差で表現した場合の決定係数は，次のようになる。

$$R^2 = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{\mathbf{y}'\mathbf{y}} \tag{34}$$

3．有意性検定および信頼区間

ここまでくると，$\hat{\beta}i$ の信頼区間および有意検定を取り扱うことができる。このために，これまでの4つの仮定に加えて新たに，次の仮定

$$u_t \sim N \tag{4－e}$$

を追加する[2]。従って，（4－a）（4－b）および（4－e）は，簡潔に

$$\mathbf{u} \sim N(\mathbf{0},\ \sigma^2\mathbf{I}_n) \tag{35}$$

と表現することができる。正規分布に従うu_i $(i=1,2\cdots k)$ の確率密度関数は,

$$f(u_i) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} exp \quad - \frac{(u_i-0)^2}{2\sigma^2} \qquad i=1,2,\cdots,\ k$$

であり, u_i, u_j, $i \neq j$ は互いにそれぞれ独立であるから, 多変量正規分布の同時確率密度関数$f(u_1,\ u_2,\ \cdots,\ u_n)$ は,

$$f(u_1, u_2, \cdots, u_n) = \frac{1}{(2\pi\sigma^2)^{n/2}} e\ x\ p - \frac{1}{2\sigma^2} [\sum_{i}^{n}(u_i-0)^2]$$

$$i=1,2,\cdots,\ n$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e\ x\ p \left\{ - \frac{(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{2\sigma^2} \right\} \tag{36}$$

となる。新たに正規分布に従うという仮定が追加されたから, 母集団の同時確率密度関数は確定する。

従って, ここで標本からの推定は最大尤度法を利用できる。すなわち, 目前の標本は, $\boldsymbol{\beta}$ に関して尤度の一番大きい母集団から抽出されたと考える。技術的には, これは$f(u_1, u_2, \cdots, u_n)$ をβに関して極大にすればよい。いいかえれば, $(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$ を極小にすることと同等である。すなわち,

$$\frac{\partial}{\partial \hat{\boldsymbol{\beta}}} f(u_1, u_2, \cdots, u_n) = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0}$$

$$\therefore \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} \tag{7}$$

すなわち $\hat{\boldsymbol{\beta}}$ の最尤推定量は, 最小二乗推定量と同じものになる。$\hat{\boldsymbol{\beta}}$ は最小二乗推定量の場合と同様に, $\boldsymbol{\beta}$ と多変量正規分布に従う $\mathbf{u}$ の線型関数との和で表わされる。従って, 最尤推定量 $\hat{\boldsymbol{\beta}}$ の期待値, 分散はそれぞれ (9) 式, (10) 式および, (4 − a) 式を考慮することによって

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta},\ \sigma^2(\mathbf{X}'\mathbf{X})^{-1}) \tag{37}$$

と表わせる。最尤推定量は, 不偏性を除く望ましい基準である一致性, 有効性, 十分性をここでは証明をしないが保障されている。この場合, 明らかに

2) u_i ($i=1,2\cdots k$) が正規分布に従わない場合であっても, 中心極限定理によってそれら k 個の変数の平均は, 漸近的に正規分布に近づくことが証明されている。

不偏性も満たされている。(37) 式が成り立てば，標準正規分布表を使って $\boldsymbol{\beta}$ に対する有意性検定と信頼区間を求めることができる。

まず，有意性検定は

$$Z=\frac{\hat{\beta}_i-\beta_i}{\sigma\sqrt{a_{ii}}} \qquad i=1,2,\cdots,k \tag{38}$$

で示されているように偏差$\hat{\beta}_i-\beta_i$ を$\hat{\beta}$の標準偏差で基準化して新しい変数 Z に変数変換をすると，Zは標準正規分布に従うことから行える。a_{ii} は $(\mathbf{X}'\mathbf{X})^{-1}$の第 i 行第 i 列の要素を表す。従って，βに対する信頼係数 $100(1-\varepsilon)$ %の信頼区間は，

$$\hat{\beta}_i \pm Z_{\varepsilon/2}\sqrt{a_{ii}}\ \sigma \tag{39}$$

で求められる。

しかし，もし分散σ^2 が未知の場合には，有意性検定および信頼区間は，t 分布を使って求められる。t 分布は，ゴセットが "Student" のペンネームで1908年に発表した論文にさかのぼる。ここから，別名Student 分布ともいう。t 分布は，正規分布とχ^2 分布の比から導かれる。

(20) 式において$\mathbf{e}'\mathbf{e}$は u の 2 次形式で以下のように示された。

$$\mathbf{e}'\mathbf{e}=\mathbf{u}'\mathbf{M}\mathbf{u}=\mathbf{u}'[\mathbf{I}_n-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}]\mathbf{u} \tag{20}$$

(20) 式の$\mathbf{M}$は対称かつ巾等行列であり，そのトレースは $n-k$ であった。従って$\mathbf{P}'\mathbf{M}\mathbf{P}=\mathbf{E}_{n-k}$となるような直交行列$\mathbf{P}$が必らず存在する。この直交行列$P$を使って $\mathbf{u}$ ベクトルを $\mathbf{v}$ ベクトルに変換する。すなわち，

$$\mathbf{u}=\mathbf{P}\mathbf{v} \quad or \quad \mathbf{v}=\mathbf{P}'\mathbf{u} \tag{40}$$
$$(\because \mathbf{P}^{-1}=\mathbf{P}')$$

(40) 式を (20) 式に代入すると，

$$\begin{aligned}\mathbf{e}'\mathbf{e}&=\mathbf{u}'\mathbf{M}\mathbf{u}=\mathbf{v}'\mathbf{P}'\mathbf{M}\mathbf{P}\mathbf{v}=\mathbf{v}'\mathbf{E}_{n-k}\mathbf{v}\\&=v_1^2+v_2^2+\cdots v_{n-k}^2\end{aligned} \tag{41}$$

(41) 式は平均 0，分散σ^2に従った $n-k$ 個の変数の残差平方の合計を表わしている。

従って，$\dfrac{\mathbf{e}'\mathbf{e}}{\sigma^2}$ は自由度$n-k$のχ^2分布に従う。残る問題点は$\mathbf{e}'\mathbf{e}$が$\hat{\boldsymbol{\beta}}$と独立であ

ることを示せばよい。これは $\mathbf{e}$ と $\hat{\boldsymbol{\beta}}$ とが独立であることを示せばよい。すなわち,

$$E[\mathbf{e}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}')] = \mathbf{0}$$

であればよい。$\mathbf{e}$ および $\hat{\boldsymbol{\beta}}$ はそれぞれ

$$\mathbf{e} = [\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{u}$$

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$$

であるから

$$E[\mathbf{e}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'] = E\{[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{u}\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\}$$

$$= \sigma^2\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} - \sigma^2\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \mathbf{0} \tag{42}$$

すなわち, $\mathbf{e}$ と $\hat{\boldsymbol{\beta}}$ はおのおの正規変量の線型関数であるから, それらは, 互いに独立に分布していることがわかる。分散 σ^2 が未知の場合, この結果を使って有意検定と β_i の信頼区間を求めることができる。(37) 式を再表現しておこう。

$$\hat{\beta}_i \sim N(\beta_i,\ a_{ii}\sigma^2) \qquad i = 1, 2, \cdots, k \tag{37$'$}$$

$\dfrac{\Sigma e^2}{\sigma^2}$ は自由度 $n-k$ の独立 χ^2 分布である。従って, t 分布の定義から,

$$t = \frac{\dfrac{\hat{\beta}_i - \beta}{\sqrt{a_{ii}\ \sigma^2}}}{\sqrt{\dfrac{1}{n-k}\ \dfrac{\Sigma e_i{}^2}{\sigma^2}}} = \frac{\hat{\beta}_i - \beta}{\sqrt{\dfrac{\Sigma e_i{}^2}{n-k}}\sqrt{a_{ii}}} \tag{43}$$

となる。分子の変数は標準正規分布 $N(0, 1)$ に従い, 分母の変数は自由度 $n-k$ の χ^2_{n-k} に従う。a_{ii} は $(\mathbf{X}'\mathbf{X})^{-1}$ の第 i 行第 i 列の主対角要素を示している。

β_i の任意の有意検定をテストするには,(43) 式に任意の値を代入して検定すればよい。

たとえば, $\beta_i = 0$ の仮設を検定する場合, (43) 式は,

$$t = \frac{\hat{\beta}_i}{\sqrt{\dfrac{\Sigma e_i{}^2}{n-k}}\ \sqrt{a_{ii}}} \tag{44}$$

となり，t 統計量を t 分布表からの値と比較して検定すればよい。また，β_i に対する信頼係数100(1 − ε) %の信頼区間は，

$$\hat{\beta}_i \pm t_{\varepsilon/2}\sqrt{\frac{\Sigma e_i}{n-k}}\sqrt{a_{ii}} \tag{45}$$

となる。

さらに，β_iのいくつかあるいは全部を同時に検定するには，これまでの理論を発展させることが必要である。これには偏差の形を使う。まず $k-1$ 個の新しい変数$z_2\cdots\cdots z_k$を，偏差$x_2\cdots\cdots x_k$を使って，次のように定義する。

$$\begin{aligned} z_{2i} &= w_{22}x_{2i} \\ z_{3i} &= w_{32}x_{2i} + w_{33}x_{3i} \qquad\qquad i=1,2,\cdots,\ n \\ &\ \vdots \\ z_{ki} &= w_{k2}x_{2i} + \cdots\cdots + w_{kk}x_{ki} \end{aligned} \tag{46}$$

ここで w は次の条件を満足するようにつくられる。

$$\left.\begin{aligned} &\sum_{i=1}^{n} z_{ji}^2 = 1 \quad i=2,3\cdots\cdots k \\ &\sum_{i=1}^{n} z_{ji}\, z_{li} = 0 \quad j,l=2,\cdots\cdots k : j \neq l \end{aligned}\right\} \tag{47}$$

すなわち，Zは直交変数である。(47) の条件式は，行列で表わすと，

$$\mathbf{Z}'\mathbf{Z} = \mathbf{I}_{k-1} \tag{47$'$}$$

また，(46) 式を行列で表わすと，

$$\mathbf{Z} = \mathbf{X}\mathbf{W} \tag{46$'$}$$

すなわち，

$$\begin{pmatrix} z_{21} & z_{31} & \cdots & z_{k1} \\ z_{22} & z_{32} & \cdots & z_{k2} \\ \vdots & \vdots & & \vdots \\ z_{2n} & z_{3n} & \cdots & z_{kn} \end{pmatrix} = \begin{pmatrix} x_{21} & x_{31} & \cdots & x_{k1} \\ x_{22} & x_{32} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{2n} & x_{3n} & \cdots & x_{kn} \end{pmatrix} \begin{pmatrix} w_{32} & w_{32} & \cdots & w_{k2} \\ 0 & w_{33} & \cdots & w_{k3} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_{kk} \end{pmatrix} \tag{46$''$}$$

w は (46)″ 式から次のようにして計算される。

まず

$$w_{22}^2 = 1 / \Sigma x_2^2$$

より，条件$\Sigma z_2^2 = 1$ を使って，w_{22}が求められる。

次のw_{32}とw_{33}は

$$w_{32}\Sigma x_2^2 + w_{33}\Sigma x_2 x_3 = 0$$

$$w_{32}^2\Sigma x_2{}^2 + w_{33}^2\Sigma x_3{}^2 + 2\,w_{32}\,w_{33}\Sigma x_2 x_3 = 1$$

より，条件$\Sigma z_2 z_3 = 0$，$\Sigma z_3^2 = 1$を使って求められる。すなわち，

$$w_{33}^2 = \Sigma x_2{}^2 / \left[\Sigma x_2{}^2 \Sigma x_3{}^2 - (\Sigma x_2 x_3)^2\right]$$

$$w_{32} = -w_{33}\Sigma x_2 x_3 / \Sigma x_2{}^2$$

と求められる。このようにして，wの値は次々に求めることができる。wの集まりの値は唯一つではないがX変数の間に完全相関がない限り，存在する。一般に（47)′式と（46)′式より

$$\mathbf{Z'Z} = \mathbf{W'X'XW} = \mathbf{I}$$

さて，$\mathbf{X}$のかわりに(46)′から導いた$\mathbf{X} = \mathbf{ZW}^{-1}$を（32）式に代入すると，

$$\mathbf{y} = \mathbf{ZW}^{-1}\boldsymbol{\beta} + \mathbf{u} - \bar{\mathbf{u}} = \mathbf{ZW}^{-1}\hat{\boldsymbol{\beta}} + \mathbf{e} \qquad (48)$$

$\boldsymbol{\beta}^* = \mathbf{W}^{-1}\boldsymbol{\beta}$，$\hat{\boldsymbol{\beta}}^* = \mathbf{W}^{-1}\hat{\boldsymbol{\beta}}$ と定義すると，（48）式は

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\beta}^* + \mathbf{u} - \bar{\mathbf{u}} = \mathbf{Z}\hat{\boldsymbol{\beta}}^* + \mathbf{e} \qquad (49)$$

となる。

（49）から，$\mathbf{y}$は係数$\boldsymbol{\beta}^*$をともなった直交変数$\mathbf{Z}$の線型関数として表現されている。そして攪乱項あるいは$\mathbf{y}$は二者択一的に係数$\hat{\boldsymbol{\beta}}^*$をともなった直交変数の線型関数となっている。$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$がわかっているから，

$$\hat{\boldsymbol{\beta}}^* = \mathbf{W}^{-1}(\mathbf{X'X})^{-1}\mathbf{X'y} \qquad (\because \hat{\boldsymbol{\beta}}^* = \mathbf{W}^{-1}\hat{\boldsymbol{\beta}})$$

$$= \mathbf{W}^{-1}(\mathbf{X'X})^{-1}(\mathbf{W}^{-1})\,\mathbf{Z'y} \qquad (\because \mathbf{Z} = \mathbf{XW}) \qquad (50)$$

しかし，$\mathbf{Z} = \mathbf{XW}$から

$$\mathbf{Z'Z} = \mathbf{W'X'XW}$$

$$\therefore\ (\mathbf{Z'Z})^{-1} = \mathbf{W}^{-1}(\mathbf{X'X})^{-1}(\mathbf{W'})^{-1}$$

$$= \mathbf{W}^{-1}(\mathbf{X'X})^{-1}(\mathbf{W}^{-1})'$$

従って，（50）式は次のようになる。

$$\therefore\ \hat{\boldsymbol{\beta}}^* = (\mathbf{Z'Z})^{-1}\mathbf{Z'y} \qquad (51)$$

すなわちこれは，（49）式の最小二乗推定量である。ここで(47)′と，それから導かれる（$\mathbf{Z'Z})^{-1} = \mathbf{I}_{k-1}$を考慮すると，（51）式は，さらに

$$\hat{\boldsymbol{\beta}}^* = \mathbf{Z}'\mathbf{y} \tag{52}$$

という単純な式になる。すなわち（52）式は，

$$\hat{\beta}_2^* = \sum_{i=1}^{n} z_{2i} y_i$$

$$\hat{\beta}_3^* = \sum_{i=1}^{n} z_{3i} y_i$$

$$\vdots$$

$$\hat{\beta}_k^* = \sum_{i=1}^{n} z_{ki} y_i$$

である。

（10）式からもわかるように，（51）式の$\hat{\boldsymbol{\beta}}^*$の分散は，

$$Var(\hat{\boldsymbol{\beta}}^*) = \sigma^2 (\mathbf{Z}'\mathbf{Z})^{-1} = \sigma^2 \mathbf{I}_{k-1} \tag{53}$$

かくして，$\hat{\beta}_i^*$はそれぞれ平均β_i^*，分散σ^2をもった独立正規分布に従っている。（53）式より

$$\frac{\hat{\beta}_i^* - \beta_i^*}{\sigma} \sim N(0, 1)$$

と基準化すると独立という性質から，

$$\frac{\sum_{i=2}^{k} (\hat{\beta}_i^* - \beta_i^*)^2}{\sigma^2} \sim \chi^2_{k-1}$$

また，

$$\frac{\sum_{i=1}^{n} e_i^2}{\sigma^2} \sim \chi^2_{n-k}$$

であったから，これらの比を作れば，F分布に従う新しい変数F統計量ができる。

$$F = \frac{\sum_{i=2}^{k} (\hat{\beta}_i^* - \beta_i^*)^2 / (k-1)}{\sum_{i=1}^{n} e_i^2 / (n-k)} \tag{54}$$

すなわち (54) 式のFは，自由度$k-1$および$n-k$のF分布に従う。

いま，$\hat{\beta}_2^* = \beta_3^* = \cdots\cdots = \beta_k^* = 0$という仮設を考えてみよう。$\boldsymbol{\beta} = \mathbf{W}\boldsymbol{\beta}^*$であるから，この仮説は，$\beta_2 = \beta_3 = \cdots\cdots = \beta_k = 0$という仮説と同じものになる。従って，$\boldsymbol{\beta}^* = \mathbf{0}$仮説はすべての係数の同時検定である。すなわち，$k-1$個の変数$X_2, X_3 \cdots\cdots X_k$が$Y$に何らかの影響するかどうかの検定である。

(54) より

$$F = \frac{\sum_{i=2}^{k} \hat{\beta}_i^{*2} / (k-1)}{\sum_{i=1}^{n} e_i^2 / (n-k)} \tag{55}$$

ここで$\sum_{i=2}^{k} \hat{\beta}_i^{*2}$は次のようになる。

$$\begin{aligned}
\sum_{i=2}^{k} \hat{\beta}_i^{*2} &= \hat{\boldsymbol{\beta}}^{*\prime} \hat{\boldsymbol{\beta}}^* = (\mathbf{W}^{-1}\hat{\boldsymbol{\beta}})' \mathbf{Z}' \mathbf{y} \quad (\because \hat{\boldsymbol{\beta}}^* = \mathbf{W}^{-1}\hat{\boldsymbol{\beta}} = \mathbf{Z}'\mathbf{y}) \\
&= \hat{\boldsymbol{\beta}}' (\mathbf{W}^{-1})' \mathbf{W}' \mathbf{X}' \mathbf{y} \quad (\because \mathbf{Z} = \mathbf{X}\mathbf{W}) \\
&= \hat{\boldsymbol{\beta}}' (\mathbf{W}')^{-1} \mathbf{W}' \mathbf{X}' \mathbf{y} \\
&= \hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{y}
\end{aligned} \tag{56}$$

さらに，ここで (34) 式より，

$$\hat{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{y} = \mathbf{y}'\mathbf{y} \cdot R^2$$

従って (24) 式の$\mathbf{e}'\mathbf{e}$は次のようになる。

$$\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y}(1 - R^2) \tag{57}$$

これら (56) および (57) 式を，(55) 式に代入すると，

$$F = \frac{\mathbf{y}'\mathbf{y} \cdot R^2 / (k-1)}{\mathbf{y}'\mathbf{y}(1-R^2) / (n-k)} = \frac{R^2 / (k-1)}{(1-R^2) / (n-k)} \tag{58}$$

この結果を第1表の分散分析表に要約しておこう。

第　1　表

変　動　要　因	平　　方　　和	自由度	平均平方(不偏分散)
X_2, X_3, ……, X_k	$\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}=\mathbf{y}'\mathbf{y}\cdot R^2$	$k-1$	$\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}/k-1$
残　　　差	$\mathbf{e}'\mathbf{e}=\mathbf{y}'\mathbf{y}(1+R^2)$	$n-k$	$\mathbf{e}'\mathbf{e}/n-k$
総　変　動	$\mathbf{y}'\mathbf{y}=\sum_{i=1}^{n} y_i^2$	$n-k$	

最後に，単一の係数に関するF検定について説明しておこう。

$\boldsymbol{\beta}^*=\mathbf{W}^{-1}\boldsymbol{\beta}$の$\mathbf{W}$は，上三角行列でその逆行列も上三角行列である。たとえば，

$$\mathbf{W}=\quad\mathbf{W}=\begin{pmatrix}1&2&3\\0&2&1\\0&0&1\end{pmatrix}$$

$$\mathbf{W}^{-1}=\frac{1}{2}\begin{pmatrix}2&-2&-4\\0&1&-1\\0&0&2\end{pmatrix}=\begin{pmatrix}1&-1&-2\\0&\frac{1}{2}&-\frac{1}{2}\\0&0&1\end{pmatrix}$$

従って，$\boldsymbol{\beta}^*=\mathbf{W}^{-1}\boldsymbol{\beta}$は，次のようになる。

$$\boldsymbol{\beta}^*=\begin{pmatrix}\beta_2^*\\\beta_3^*\\\beta_4^*\end{pmatrix}=\begin{pmatrix}1&-1&-2\\0&\frac{1}{2}&-\frac{1}{2}\\0&0&1\end{pmatrix}\begin{pmatrix}\beta_2\\\beta_3\\\beta_4\end{pmatrix}=\begin{pmatrix}\beta_2-\beta_3-2\beta_4\\\frac{1}{2}\beta_3-\frac{1}{2}\beta_4\\\beta_4\end{pmatrix}$$

この例からもわかるように，

$$\beta_2^*=f_2(\beta_2,\ \beta_3,\ \beta_4)$$
$$\beta_3^*=f_3(\beta_3,\ \beta_4)$$
$$\beta_4^*=f_4(\beta_4)$$

一般に，

$$\begin{aligned}\beta_2^*&=f_2(\beta_2,\ \beta_3,\ \cdots\cdots,\ \beta_k)\\\beta_3^*&=f_3(\beta_3,\ \cdots\cdots,\ \beta_k)\\&\ \vdots\\\beta_k^*&=f_k(\beta_k)\end{aligned}$$

すなわち，f_iはβ_iに関して線型同次関数であることがわかる。

よって，$\beta_k = 0$ と仮説は，$\beta_k^* = 0$ という仮説と同等である。

$\beta_k^* = 0$ 仮説を検定するには，

$$\frac{\hat{\beta}^{*2}}{\sigma^2} \sim \chi_1^2$$

$$\frac{\sum_{i=1}^{n} e_i{}^2}{\sigma^2} \sim \chi_{n-k}^2$$

の2つの比を作ると，

$$F = \frac{\hat{\beta}_k^{*2}}{\sum_{i=1}^{n} e_i{}^2 / (n-k)} \tag{59}$$

は，自由度（1，n—k）のF分布に従う。

$\sum_{i=2}^{k-1} \hat{\beta}_i^{*2}$は，$z_2$，……，$z_{k-1}$ あるいは，X_2，……，X_{k-1} によって Yの総変動が説明される説明二乗和を表わす。$\sum_{i=2}^{n} \hat{\beta}_i^{*2}$は，すべての変数$X_2$，……，$X_{k-1}$，$X_k$によって$Y$の総変動がどれだけ説明されるかを示す説明二乗和を示す。従って，$\hat{\beta}_k^{*2}$はX_kをX_2，……，X_{k-1} に加えるときの説明二乗和の増加分を表わす。これを表に要約したものが第2表の分散分析表である。

第　2　表

変動要因	平方和	自由度	平均平方
X_2，……，X_{k-1}	$\sum_{i=2}^{k-1} \hat{\beta}_i^{*2}$	$k-2$	
X_k	$\hat{\beta}_k^{*2}$	1	$\hat{\beta}_k^{*2}$
X_2，……，X_k	$\sum_{i=2}^{k} \hat{\beta}_i^{*2}$	$k-1$	
残差	$\sum_{i=1}^{n} e_i{}^2$	$n-k$	$\sum_{i=1}^{n} e_i / n-k$
総変動	$\sum_{i=1}^{n} y_i{}^2$	$n-1$	

この検定は既に論じた t 分布を使う検定と等しい。(44) 式に示されたように，$\beta_k = 0$ の仮説の検定は，

$$t=\frac{\hat{\beta}_k}{\sqrt{\sum_{i=1}^{n} e_i^2 / n-k}\ \sqrt{a_{kk}}} \tag{44)'$$

が，自由度 $n-k$ をもつ t 分布に従った。a_{kk}は，$(X'X)^{-1}$ の k 番目の主対角要素である。この t 分布を 2 乗したものをFとおくと，

$$F=\frac{\hat{\beta}_k^2}{\left(\sum_{i=1}^{n} e_i^2 / n-k\right) a_{kk}} \tag{60}$$

(59) 式の分散分析のF統計量が

$$F=\frac{\hat{\beta}_k^{*2}}{\sum_{i=1}^{n} e_i^2 / n-k} \tag{59}$$

であったから，もし

$$\hat{\beta}_k^2 = a_{kk}\hat{\beta}_k^{*2}$$

であれば (44)′ 式と (60) 式は同等であることが判明する。一方，

$$\hat{\boldsymbol{\beta}}^* = \mathbf{W}^{-1}\hat{\boldsymbol{\beta}}$$

$$\hat{\boldsymbol{\beta}} = \mathbf{W}\hat{\boldsymbol{\beta}}^*$$

より

$$\hat{\beta}_k = w_{kk}\hat{\beta}_k^*$$

従って，$w_{kk}^2 = a_{kk}$を示せばよい。

(47)′式及び (46)′式より，

$$\mathbf{Z}'\mathbf{Z} = \mathbf{W}'\mathbf{X}'\mathbf{X}\mathbf{W} = \mathbf{I}_{k-1}$$

従って，

$$(\mathbf{X}'\mathbf{X})\mathbf{W} = (\mathbf{W}')^{-1}$$

$$(\mathbf{X}'\mathbf{X}) = (\mathbf{W}')^{-1}\mathbf{W}^{-1} = (\mathbf{W}\mathbf{W}')^{-1}$$

$$\therefore (\mathbf{X}'\mathbf{X})^{-1} = \mathbf{W}\mathbf{W}'$$

すなわち

$$
\mathbf{WW}' = \begin{pmatrix} w_{22} & w_{32} & \cdots & w_{k2} \\ 0 & w_{33} & \cdots & w_{k3} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & w_{kk} \end{pmatrix} \begin{pmatrix} w_{22} & 0 & \cdots & 0 \\ w_{32} & w_{33} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ w_{k2} & w_{k3} & \cdots & w_{kk} \end{pmatrix} = \begin{pmatrix} w_{22}^2 + w_{32}^2 + \cdots + w_{k2}^2, & \cdots, & w_{k2}w_{kk} \\ \vdots & \ddots & \vdots \\ w_{kk}w_{k2}, & \cdots\cdots, & w_{kk}^2 \end{pmatrix}
$$

$$
w_{kk}^2 = a_{kk}
$$

かくのごとく $(\mathbf{X}'\mathbf{X})^{-1}$の第 k 番目の主対角要素はw_{kk}^2で(38)式あるいは, (43)式のa_{kk}に該当することが証明される。すなわち伝統的な回帰係数β_iに関する t 検定は, 変数X_2, X_3, …, X_{i-1}, X_{i+1},…, X_k に X_i 新たに加えることによって生ずる説明される平方和の増分が残差平方和に比べて有意に大きいかどうかを決めることである。

第 2 表に示した分散分析表は, 1 変数からいくつかの変数のグループの検定に拡張できる。たとえば$X_{r+1}\cdots X_k$を加えることにしよう。$\hat{\beta}_i^*$は分散σ^2で平均β_i^*のまわりにそれぞれ独立に正規分布するから, 次の 2 つの比

$$
F = \frac{\sum_{i=r+1}^{k} (\hat{\beta}_i^* - \beta_i^*)^2 \diagup (k-r)}{\sum_{i=1}^{n} e_i{}^2 \diagup (n-k)} \tag{61}
$$

は, 自由度 ($k-r$, $n-k$) のF分布に従う。たとえば, $\beta_{r+1} = \cdots = \beta_k = 0$ の仮説は,

$$
\beta_{r+1}^* = \cdots\cdots = \beta_k^* = 0
$$

と同等であった。従って(61)式は

$$
F = \frac{\sum_{i=1}^{k} \hat{\beta}_i^{*2} \diagup (k-r)}{\sum_{i=1}^{n} e_i^2 \diagup (n-k)} \tag{62}
$$

と書き換えられる。この場合の分散分析表は次のようになる。

第　3　表

変数要因	平方和	自由度	平方平均(不偏分数)
$X_2, X_3, \cdots\cdots, X_r$	$\sum_{i=2}^{r}\hat{\beta}_i^{*2}$	$r-1$	
$X_{r+1}, \cdots\cdots, X_k$	$\sum_{i=r+1}^{k}\hat{\beta}_i^{*2}$	$k-r$	$\sum_{i=r+1}^{k}\hat{\beta}_i^{*2}/(k-r)$
$X_2, \cdots\cdots\cdots, X_k$	$\sum_{i=2}^{k}\hat{\beta}_i^{*2}$	$k-1$	
残　差	$\sum_{i=1}^{n}e_i^2$	$n-k$	$\sum_{i=1}^{n}e_i^2/(n-k)$
総　変　動	$\sum_{i=1}^{n}y_i^2$	$n-1$	

4. あ と が き

以上で一般の線型回帰モデルの推定および検定の理論について論じてきた。しかし，本論で示されたこれらの方法を実際に適用する場合には，基本的仮定を前提としていることを忘れてはならない。もし，これら4つの基本的仮定の1つでも満足しなければ，モデルそのものを修正しなければならない。たとえば攪乱項u_iが互いに独立でなければ，いいかえれば自己相関がある場合はそれを考慮しなければならないし，また分散が一定でない不等分散の時には，モデルは一般化最小二乗法によって推定されなければならない。

また，任意の2つの説明変数X_i, X_jが互いに相関関係にある場合には多重共線性 (Multicollinearity) が問題が起り，２段階最小二乗法によって推定しなければならない。さらに，説明変数が遅れを伴った時にはラグ (lag) の問題が起ったり，母集団に異質な構造の併存が予想される場合にはダミー変数を使ったりする。いずれにしても，いろいろなケースに応じた推定法が適用されている。

しかし，いかなるケースにも適用できるような推定法はなく，個々のケー

スに応じたものしか開発されていない。医学の治療と同じようにあらゆる病気に万べんなくきく万能薬はなく，１つ１つの問題に対応していく推定法である。その意味では，回帰分析の推定法は発展途上にあると言ってよいであろう。

上記にあげた推定法やその他の推定法については次の機会に発表，あるいは今後の課題として本論を終えることにする。

（本論は，記号や説明の進め方について*Johnston*〔20〕の文献に従い執筆したことを付記しておきます。）

参考文献

〔１〕 内田忠夫他編集 『近代経済学講義』 計量分析篇１ 有斐閣1968

〔２〕 竹内 啓著 『計量経済学の研究』 東洋経済1972

〔３〕 上野裕也・村上泰亮編 『日本経済の計量分析』 岩波書店1975

〔４〕 山田 勇編 『計量経済学講義』 青林書院1972

〔５〕 内田忠夫・福地崇生著 『計量経済学』 筑摩書房1975

〔６〕 森口親司著 『計量経済学』 岩波書店1974

〔７〕 経済企画庁経済研究所編 『短期経済予測マスターモデルの研究』 研究シリーズ21号 大蔵書印刷1970

〔８〕 建元正弘・市村真一編 『日本経済の計量分析』 東洋経済1970

〔９〕 浜田文雅著 『設備投資行動の計量分析』 東洋経済1971

〔10〕 佐和隆光著 『計量経済学の基礎』 東洋経済1970

〔11〕 金森久雄編 『日本経済の変動と予測』 日本経済新聞社1969

〔12〕 森田優三著 『経済統計読本』 東洋経済1970

〔13〕 竹内　啓著　『数理統計学』　東洋経済1963

〔14〕 岩田暁一著　『経済分析のための統計的方法』　東洋経済1967

〔15〕 守谷栄一監修 井口晴弘著　『多変量解析とコンピュータプログラム』　日刊工業1972

〔16〕 奥野忠一他著　『多変量解析法』　日科技連1971

〔17〕 建元正弘 真継　隆 著　『社会人のための計量経済学』　日本経済新聞社1973

〔18〕 福地崇生著　『計量経済学』　東洋経済1962

〔19〕 宮川公男著　『計量経済学入門』　日本経済新聞社1966

〔20〕 J.Johnston, econometric methods, 2nd edition, McGraw-Hill 1972

〔21〕 L. R. Klein, A textbook of econometrics, Row Peterson Maruzen 1953

〔22〕 M. K. Evans, 『マクロ経済活動の分析』上・下 片渕泰訳　鹿島出版会1972

〔23〕 J. L. Bridge, Applied Econometrics, North-Holland Pub. co., 1971

〔24〕 E. Kuh R. Schmalensce 『マクロ経済モデル』 浜田文雅訳 マグロウヒル好学社1975

〔25〕 A. S. Goldberger 『計量経済学の理論』 福地崇生 森口親司 共訳 東洋経済1970

〔26〕 S. S. Wilks, Mathematical Statistics, 2nd edition, WILEY TOPPAN 1962

〔27〕 J. Kmenta, Elements of Econometrics, Macmillan, 1971

〔28〕 H. Theil J. C. G. Boot T. Kloek 『ORと計量経済学』 関根智明訳 好学社1968

〔29〕 T. Yamane, Statistics, 3rd edition, Harper & Row, 1973